MEME16203Linear Models

Assignment 2

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty:	FES	Unit Code:	MEME16203
Course:	MAC	Unit Title:	Linear Models
Year:	1,2	Lecturer:	Dr Wong Wai Kuan
Session:	May 2024		
Due by:	28/6/2024		

Q1. The following experiment was performed to determine if treatment with a substance called MDL is effective in controlling blood pressure. Sixteen rabbits were used in this experiment. The rabbits were randomly divided into two groups with eight rabbits in each group. The rabbits in one group were treated with MDL and the rabbits in the other group were given a plcebo. the rabbits were then randomly assigned to receive intravenous injection of one of four levles ($x_1 = 25, x_2 = 50, x_3 = 75$, or $x_4 = 100$ ug/kg) of phenylbiguanide (PBG) which is known to increase blood pressure. In this way, two rabbits were randomly assigned to each of the eight combinations of the two factors. The measured response was the percent increase in blood pressure relative to a measurement taken on each rabbit before this experiment was begun.
The data are as follows:

	Level of PBG (ug/kg)			
	25	50	75	100
Treatment with MDL	y_{111}	y_{121}	y_{131}	y_{141}
	y_{112}	y_{122}	y_{132}	y_{142}
Placebo	y_{211}	y_{221}	y_{231}	y_{241}
	y_{221}	y_{222}	y_{232}	y_{242}

Consider a Gauss-Markov model

$$y_{ijk} = \mu + \alpha_i + \gamma X_j + \epsilon_{ijk} \qquad \text{(Model 1)}$$

where

- y_{ijk} is the percent increase in blood pressure for the k^{th} rabbit assigned to the j^{th} level of the PBG factor and the i^{th} level of the MDL factor,
- X_j denote the level of PBG administered to the rabbit,
- ϵ_{ijk} denotes a random error with $\epsilon_{ijk} \sim NID(0, \sigma^2)$ where $\sigma^2 > 0$, and
- $\boldsymbol{\beta} = (\mu, \alpha_1, \alpha_2, \gamma)^T$ is a vector of unknown non-random parameters.

Use this model to answer the following questions. (You may express your answers in matrix notation, but define any new notation that you introduce.)

(a) Model (1) can be expressed in the form $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. Do not impose any restriction on the parameters, write down matrix of $\mathbf{X}$ in kronecker form.

(b) Show that $\mu + \alpha_1 + \gamma x$ is estimable for any $x \in \mathbf{R}$.

(c) Show that $\mathbf{X}$ in part (a) has the same column space as

$$\mathbf{W} = \begin{bmatrix} 1 & -1 & -3 \\ 1 & -1 & -3 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 3 \\ 1 & -1 & 3 \\ 1 & 1 & -3 \\ 1 & 1 & -3 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & 1 & 3 \end{bmatrix}$$

(d) Determine the ordinary least squares estimator of $\mu + \alpha_1 + \gamma x$ for any level of the PBG in terms of y_{ijk} and their means.

(25 marks)

Q2. The response time in milliseconds was determined for three different types of circuits that could be used in an automatic valve shutoff mechanism. The results are shown in the following table.

Circuits Type	Response Time
1	9 12 10
2	20 21 23
3	6 5 8

Consider the linear model

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \text{ for } i = 1,2,3 \text{ and } j = 1,2,3$$

where where μ is the overall mean, τ_i is the circuit type content effects and $\epsilon_{ij} \sim N(0, \sigma^2)$ is the random error.

(a) Write model above in the form $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$. Do not impose any restriction on the parameters.

(b) Obtain one of the generalized inverse of $\mathbf{X^T X}$, $\mathbf{G}$.

(c) Use the generalized inverse you obtained in part(b) to compute a solution to the normal equations, $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\tau}_3 \end{bmatrix}$.

(d) Using your solution $\hat{\boldsymbol{\beta}}$ to the normal equation from part (c), estimates $\tau_1 - 2\tau_2 + \tau_3$.

(e) Is your estimate in part(d) the unique BLUE? Explain.

(25 marks)

Q3. Data were collected to study the effect of temperature on the yield of a chemical process. Two different catalyst A and B, were used in the study. Yields were measured under 4 different temperatures for each catalyst. The data are as follows:

Run	Yield (Grams), y	Temperatures (°C), T	Catalyst
3	26	80	A
10	28	85	A
4	15	90	A
8	39	95	A
5	17	100	A
9	47	80	B
2	32	85	B
6	28	90	B
1	6	95	B
7	24	100	B

Each run can be considered as an independent observation. The order in which the runs were made was randomized. Consider the linear model

$$y_{ij} = \mu + \alpha_i + \beta(T_{ij} - 90) + \epsilon_{ij}, \text{ for } i = 1, 2 \text{ and } j = 1, 2, \ldots, 5$$

where

- y_{ij} is the observed yield for the run using the i^{th} catalyst and the j^{th} temperature level.
- α_i corresponds to the i^{th} catalyst
- T_{ij} is the temperature under which the process was run.

(a) Write model above in the form $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$. Do not impose any restriction on the parameters.

(b) Determine which, if any, of the following quantities are estimable. For each estimable quantity, report the value of a vector a such that $\mathbf{a^T y}$ satisfies the definition of an estimable function of $\mathbf{b}$.

(i)μ

(ii)$\mu + \alpha_2$

(iii)$\alpha_1 - \alpha_2$

(iv)$\mu + \beta T$ where T is any specified temperature

(v)$\mu + \alpha_1 + \beta T$ where T is any specified temperature

(c) Compute one of the generalized inverse of $(\mathbf{X^T X})^-$, G. Use the generalized inverse to compute a solution to the normal equations, $\hat{\boldsymbol{\beta}} = (\hat{\mu}\ \hat{\alpha}_1\ \hat{\alpha}_2\ \hat{\beta})^T$.

(d) Using your solution $\hat{\beta}$ to the normal equations from part (c), estimates $\mu + \alpha_1 + 5\beta$.

(e) Using your solution $\hat{\beta}$ to the normal equations from part (c), estimates of the lines that describe how the mean yield changes with changes in temperature are

$$\hat{y}_{1j} = \hat{\mu} + \hat{\alpha}_1 + \hat{\beta} T_{1j}$$

when catalyst A is used, and

$$\hat{y}_{2j} = \hat{\mu} + \hat{\alpha}_2 + \hat{\beta} T_{2j}$$

when catalyst B is used

Would these estimates change if you used a different solution to the normal equations? Explain.

(25 marks)

Q4. Suppose $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where for $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)'$, $E(\boldsymbol{\epsilon}) = \mathbf{0}$. A particular experiment produces $n = 5$ data points as per

x_i	21	30	26	36	88
y_i	316	392	114	139	89

Suppose that $V(\boldsymbol{\epsilon}) = \sigma^2 diag\left(\frac{1}{9}, \frac{1}{16}, \frac{1}{49}, \frac{1}{81}, \frac{1}{121}\right)$. Evaluate an appropriate BLUE of $\boldsymbol{\beta} = \begin{bmatrix}\beta_0 \\ \beta_1\end{bmatrix}$ under the model assumptions.

(25 marks)